

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,000

Open access books available

125,000

International authors and editors

140M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



The Borel-Cantelli Lemmas, and Their Relationship to Limit Superior and Limit Inferior of Sets (or, Can a Monkey Really Type Hamlet?)

Anant P. Godbole

Abstract

The purpose of this chapter is to show that if a monkey types infinitely, Shakespeare's *Hamlet* and any other works one may wish to add to the list will *each* be typed, not once, not twice, but *infinitely often* with a probability of 1. This dramatic fact is a simple consequence of the *Borel-Cantelli lemma* and will come as no surprise to anyone who has taken a graduate-level course in Probability. The proof of this result, however, is quite accessible to anyone who has but a rudimentary understanding of the concept of independence, together with the notion of limit superior and limit inferior of a sequence of sets.

Keywords: Borel-Cantelli lemma, limit superior of sets, limit inferior of sets, independence, limit theorems of probability

1. Introduction

Consider a monkey named Sue who is given a word processor with N symbols. We shall assume that these symbols include the 26 letters of the English alphabet (upper and lower case), all the Greek letters, the numbers 0 through 9, a blank space, all the standard punctuation marks (.,; – etc.), and mathematical symbols (∞ , \int , \Rightarrow , ∇ , etc.); imagine, in fact, that N is so large that the keyboard is capable of typing just anything we might fancy, in any language. (A L^AT_EX editor could do much of that too, but not in all languages!)

If Sue is handed such a machine and pounds away, randomly, it is clear that most of what she types will be complete gibberish. A stray segment of sense such as “dog eat HAY” or even “Ronnie ReAGAn our 40th PresiD ϵ η τ of USA” would surprise no one, but how often, we ask, would she successfully manage to type the Constitution of the United States, or Shakespeare's *Hamlet*, or the fundamental mathematical works of the 2026 Fields Medalist(s)?

The purpose of this chapter is to show that if Sue types infinitely, the above works (and any others that one may choose to add to the list) will *each* be typed, not once, not twice, but *infinitely often* with a probability of 1. This dramatic fact is a

simple consequence of the *Borel-Cantelli lemma* and will come as no surprise to anyone who has taken a serious graduate-level course in Probability. The proof of this result, however, is quite accessible to anyone who has but a rudimentary understanding of the concept of independence.

The reader is invited, while reading this chapter, to let his/her imagination run wild, and concoct a plethora of similar examples. A somewhat mundane objection may be raised immediately: how can Sue (or anyone else for that matter) type indefinitely? We shall not dwell on this nonmathematical problem, but will remark instead (and prove a little later) that Sue's never-ending assignment is mathematically equivalent to the task of randomly selecting a number from the interval $[0, 1]$.

We would like to mention that our problem is related to the famous "Problem of a printed line," a popular account of which can be found in George Gamow's classic book [1]. The solution presented there, however, is entirely deterministic and of a finite character: the automatic printing press considered by Gamow does not print indefinitely, and the probabilities of various outcomes are not calculated.

2. Limit superior and limit inferior of a sequence of sets

Consider any sequence $\{A_n\}_{n=1}^{\infty}$ of subsets of a set Ω . Points of Ω will be denoted by ω . We know that

$$\bigcup_{n=1}^{\infty} A_n = \{\omega \in \Omega : \omega \in A_n \text{ for some } n\}$$

and

$$\bigcap_{n=1}^{\infty} A_n = \{\omega \in \Omega : \omega \in A_n \text{ for each } n\}$$

It follows that

$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \{\omega \in \Omega : \forall n \exists k \geq n \text{ with } \omega \in A_k\}$$

To better understand this somewhat complicated set, we first let $n = 1$ and note that $\omega \in A_k$ for some $k \geq 1$, say $k = k_1$. Letting $n = k_1$, we see that ω must belong to some A_{k_2} , where $k_2 \geq k_1$. Continuing in this fashion, we see that $\omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ if and only if $\omega \in A_k$ for infinitely many k 's.

The set $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ is called the *limit superior* of the sequence $\{A_n\}_{n=1}^{\infty}$, and is denoted by $\limsup A_n$, or, $\overline{\lim} A_n$, or, rather appropriately, by $\{A_n \text{ i.o.}\}$, where i.o. stands for "infinitely often."

In a similar fashion, we observe that the set

$$\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \{\omega \in \Omega : \exists n \text{ such that } \omega \in A_k \text{ } \forall k \geq n\}$$

is a collection of those points ω that belong to *all but a finite number* of the A_n 's. $\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$ is called the *limit inferior* of the sequence $\{A_n\}_{n=1}^{\infty}$ and is usually denoted by $\liminf A_n$ or $\underline{\lim} A_n$. We prefer the notation $\{A_n \text{ a.b.f.o.}\}$ (a.b.f.o. means "all but finitely often"). Elementary symbol manipulation may be used to prove that $\underline{\lim} A_n \subset \overline{\lim} A_n$. It is easier to note, however, that if ω belongs to all but finitely many A_n 's it must necessarily belong to an infinite number of them. The above fact is just one of the many similarities between lim sups and lim infs of sets, on the one

hand, and of real numbers, on the other (recall that $\liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n$). Likewise $\liminf_{n \rightarrow \infty} a_n$ and $\limsup_{n \rightarrow \infty} a_n$ must both exist, that is,

$$-\infty \leq \liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n \leq \infty,$$

as must $\underline{\lim} A_n$ and $\overline{\lim} A_n$, with

$$\phi \subset \underline{\lim} A_n \subset \overline{\lim} A_n \subset \Omega.$$

If $\{A_n \text{ a.b.f.o}\} = \{A_n \text{ i.o.}\}$, the sequence $\{A_n\}$ is said to have a limit, which we define to be the common value of $\{A_n \text{ a.b.f.o}\}$ and $\{A_n \text{ i.o.}\}$ and denote by $\lim_{n \rightarrow \infty} A_n$ (note, again, the analogy with real sequences).

A useful dual relation between these two sets is

$$(\overline{\lim} A_n)^c = \underline{\lim} A_n^c,$$

and

$$(\underline{\lim} A_n)^c = \overline{\lim} A_n^c$$

where A^c denotes the complement of the set A . Here is an informal proof of the first of these two facts: $\omega \in \underline{\lim} A_n^c$ iff ω belongs to all but finitely many A_n^c 's iff $\omega \in$ just a finite number of the A_n 's iff $\omega \notin \overline{\lim} A_n$ iff $\omega \in (\overline{\lim} A_n)^c$.

A few examples should help familiarize the reader with the above notions: the second and the third are taken from [2]:

Example 1. If $A_1 \subset A_2 \subset \dots$, then $\overline{\lim} A_n = \underline{\lim} A_n = \cup_{n=1}^{\infty} A_n$. Likewise, if $A_1 \supset A_2 \supset A_3 \supset \dots$, then $\overline{\lim} A_n = \underline{\lim} A_n = \cap_{n=1}^{\infty} A_n$.

Example 2. If $A_{2n-1} = (-\frac{1}{n}, 1]$ and $A_{2n} = (-1, \frac{1}{n}]$, then $\underline{\lim} A_n = \{0\}$ and $\overline{\lim} A_n = (-1, 1]$.

Example 3. if A_n is the unit circle with center at $(\frac{(-1)^n}{n}, 0)$, then

$$\underline{\lim} A_n = \{(x, y) : x^2 + y^2 < 1\}$$

and

$$\overline{\lim} A_n = \{(x, y) : x^2 + y^2 \leq 1\} \setminus \{(0, 1), (0, -1)\}.$$

In what follows, the set Ω will be taken to be the *sample space* (or set of possible realizations) of a random experiment (one whose outcome cannot be predicted in advance). We shall assume that each subset of Ω that we encounter is *measurable*. In other words, each set A will be assumed to belong to the *sigma algebra* \mathcal{A} of *events*, which is a class of subsets of Ω satisfying the conditions

- $\phi \in \mathcal{A}$
- If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ and
- If $A_1, A_2, \dots \in \mathcal{A}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$.

This restriction ensures that sets such as $\overline{\lim} A_n$ and $\underline{\lim} A_n$ are themselves measurable, so that we may meaningfully talk of their probabilities $P\{A_n \text{ i.o.}\}$ and $P\{A_n \text{ a.b.f.o}\}$. We next move on to a key concept in probability:

Definition: The sequence of events $\{A_n\}_{n=1}^{\infty}$ will be said to be *independent*, if for each finite subcollection $A_{n_1}, A_{n_2}, \dots, A_{n_k}$,

$$P(A_{n_1} \cap A_{n_2} \cap \dots \cap A_{n_k}) = P(A_{n_1}) \cdot P(A_{n_2}) \cdot \dots \cdot P(A_{n_k}).$$

Stated informally, this means that the occurrence (or nonoccurrence) of any finite subcollection $\{A_{n_1}, A_{n_2}, \dots, A_{n_k}\}$ does not affect the probability of occurrence of another disjoint collection $\{A_{m_1}, A_{m_2}, \dots, A_{m_\ell}\}$.

The events $\{A_n\}_{n=1}^{\infty}$ that represent the successive outcomes of an infinite coin-tossing experiment are usually assumed, on intuitive and empirical grounds, to be independent. We shall make the same assumption regarding Sue's successive choices $\{B_n\}_{n=1}^{\infty}$ of a keyboard's key.

The *Borel-Cantelli lemma* is a two-pronged theorem, which asserts that the probability of occurrence of an infinite number of the independent events $\{A_n\}_{n=1}^{\infty}$ is zero or one:

Theorem 2.1. (*The Borel-Cantelli lemma*, [3, 4]).

- If $\{A_n\}_{n=1}^{\infty}$ is any sequence of events, then $\sum_{n=1}^{\infty} P(A_n) < \infty$ implies that $P(A_n \text{ i.o.}) = 0$.
- If $\{A_n\}_{n=1}^{\infty}$ is an independence sequence, then $P(A_n \text{ i.o.})$ equals 0 or 1 according as the series $\sum_{n=1}^{\infty} P(A_n)$ converges or diverges.

The following lemma can be proved using elementary properties of probability measures:

Lemma 2.2. If $\{A_n\}_{n=1}^{\infty}$ is an increasing (decreasing) sequence of events then $\lim_{n \rightarrow \infty} P(A_n) = P(\cup_{n=1}^{\infty} A_n) [P(\cap_{n=1}^{\infty} A_n)]$.

Proof of Theorem 2.1

- For any n , we note that

$$P(A_n \text{ i.o.}) \leq P(\cup_{k \geq n} A_k) \leq \sum_{k=n}^{\infty} P(A_k)$$

On letting $n \rightarrow \infty$, we see that

$$P(A_n \text{ i.o.}) \leq \lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} P(A_k) = 0,$$

proving part (a).

- We shall prove that $P(A_n^c \text{ a.b.f.o.}) = 0$. Let m, n ($m \geq n$) be arbitrary. Note that

$$P(\cap_{k=n}^m A_k^c) = \prod_{k=n}^m P(A_k^c) = \prod_{k=n}^m (1 - P(A_k)) \leq \exp \left\{ - \sum_{k=n}^m P(A_k) \right\},$$

where the last inequality follows from the fact that $1 - x \leq e^{-x}$ ($x \geq 0$). Lemma 2.1 now gives us that

$$\lim_{m \rightarrow \infty} P\left(\cap_{k=n}^m A_k^c\right) = P\left(\cap_{k=n}^{\infty} A_k^c\right) = 0 \quad \forall n,$$

which proves that

$$P(A_n^c \text{ a.b.f.o.}) \leq \sum_{n=1}^{\infty} P\left(\bigcap_{k=n}^{\infty} A_k^c\right) = 0,$$

proving the result.

As seen by the dates on Refs. [3, 4], the Borel-Cantelli lemmas are classical, and now part of virtually all graduate level books on Probability such as [2]. Since then, for over 100 years, the literature on the lemmas has focused on weakening the independence requirement in the second lemma, or looking at more complicated probability models that yield the same conclusions. See for example [5–7]. What distinguishes this work from these and others is that we provide a very down-to-earth application that forces the reader to come to terms with the notions of independence and infinity, as opposed to the finite samples one has in statistical situations. It is a paper that we feel can cause amusement, astonishment, false disbelief, and, ultimately, understanding. With this backdrop, we are now in a position to start establishing the claim made at the beginning of this chapter:

Corollary 2.3. *If Sue types indefinitely, by successively and independently choosing one of the N available keys, then any specific work containing a total of M characters (including all the blanks, of course) will be typed by her indefinitely often, with a probability of 1.*

Proof. Let A_1 be the event that Sue's first M random choices lead to the work being typed correctly. It is clear that $P(A_1) = \left(\frac{1}{N^M}\right)$. Similarly, let A_2 denote a successful completion of the task between the $(M+1)^{\text{st}}$ and $(2M)^{\text{th}}$ keystrokes. In general, A_n denotes the completion of a flawless job between $((n-1)M+1)^{\text{st}}$ and the $(nM)^{\text{th}}$ random strokes. It is evident that $P(A_n) = \left(\frac{1}{N^M}\right)$ ($n \geq 1$) and that the sequence $\{A_n\}_{n=1}^{\infty}$ is independent. Since $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \left(\frac{1}{N^M}\right) = \infty$, it follows by part (b) of the Borel-Cantelli lemma that $P(A_n \text{ i.o.}) = 1$. Since the probability that the work is typed correctly infinitely often is at least as large as $P(A_n \text{ i.o.})$, the proof is complete. (Notice how the events $\{A_n\}_{n=1}^{\infty}$ are defined using disjoint blocks; this guarantees their independence.)

Lemma 2.4. *If $P(C_n) = 1$ ($n = 1, 2, \dots$), then $P(\cap_{n=1}^{\infty} C_n) = 1$.*

Proof: Boole's inequality states that

$$P(\cap_{k=1}^n C_k) \geq \sum_{k=1}^n P(C_k) - (n-1).$$

Thus $P(\cap_{k=1}^n C_k) = 1$ for each n . The required conclusion is obtained on letting $n \rightarrow \infty$, and using Lemma 2.2.

Corollary 2.5. *If Sue types indefinitely, then every piece of writing (of any finite length whatsoever; published, unpublished, or yet to be written by yet unborn individuals; meaningful or gibberish) will be typed by her, infinitely often each, with a probability of 1.*

Proof. Denote the works by B_1, B_2, \dots , and set $A_n = \{B_n \text{ i.o.}\}$, ($n = 1, 2, \dots$). By Lemma 2.2, $P(A_n) = 1$ for each n and thus, by Lemma 2.4, $P(\cap_{n=1}^{\infty} A_n) = 1$, as claimed. (Note: this implies that the probability that *some* work is typed finitely often (or never) is zero.)

Example 4. (Statistical tests of hypotheses) If a fair coin is tossed infinitely often, a sequence of 10^6 consecutive heads will appear infinitely often with probability 1.

Now, if a coin (of unknown origin) were tossed a million times, and a head appeared each time, the “null” statistical hypothesis

$$H_0 : \text{The coin is fair } (p = 1/2)$$

would be summarily rejected at most conventional (5, 1, 0.00001%) levels of significance. The point to note, however, is that such “extreme” and “erratic” behavior *will* be exhibited on an infinite number of occasions by any fair coin (and by all coins with $P(H) > 0$), with a probability of 1.

Similarly, if a fair coin is tossed infinitely often, an n -long alternating sequence $HTHT \dots HT$ (n is arbitrary) will appear infinitely often, almost certainly. This fact may be compared with the conclusion of a standard nonparametric statistical procedure, *the run test*: the fair coin hypothesis would be vigorously rejected, using this test, if a large number of coin tosses yielded an alternating sequence of heads and tails.

3. A probability model for infinite coin tossing

In the above discussion, we often concluded that a particular event (e.g., *Hamlet* is typed infinitely often) occurred with probability 1. One fundamental question that we did not address, however, was the following: just what probability model describes infinite coin tossing or simian typewriting? Put another way, what are the sample spaces associated with these two experiments? And what exactly is the probability of an event defined to be? We realize then, in retrospect, that we had put the cart before the horse; various events were shown to have probability 1, by *assuming* the existence of a logically consistent probability (measure) on a sample space that had not been fully described. This practice is fairly standard in the teaching of probability; for example, sequences $\{X_n\}_{n=1}^{\infty}$ of independent and identically distributed (i.i.d.) random variables are often introduced as mathematical objects before their existence is proved, using Kolmogorov’s famous Consistency Theorem. Such an approach is often beneficial; as Billingsley [8] wrote, “It is instructive... to see the show in rehearsal as well as in performance.”

We shall start by noting that three tosses of a fair coin lead to the eight-point sample space

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

It seems reasonable to assign probability $1/8$ to each of these eight points; thus the probability $P(A)$ of *any* subset A may be defined by

$$P(A) = \frac{\text{number of points in } A}{8}$$

Our analysis is thus complete, and can easily be extended to any finite number of coin tosses. The situation gets rapidly more complicated if the coin is tossed endlessly. This experiment cannot be conceived, carried out, or justified “in practice,” and our neat conclusions would be rendered meaningless if we were unable to mathematically model our procedure. Happily, however, this is not the case. We simply let

$$\begin{aligned} \Omega &= \{\omega : \omega \text{ is an infinite sequence of H's and T's}\} \\ &= \{\omega : \omega \text{ is an infinite sequence of 1's and 0's}\} \end{aligned}$$

A typical element of Ω might be $\omega = 0010011 \dots$. It is well known (and easily proved) that Ω is an uncountable set. It seems reasonable, then, to assign probability zero to each sample point.

The next step is crucial. We identify each element of Ω with the real number in the interval $[0,1]$ that has the same binary expansion. For example, the sample outcome THTHTH ... is identified with the real number $0.01010101 \dots$ which equals

$$\frac{0}{2^1} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{1}{2^4} + \dots = \frac{1}{3}.$$

A problem arises immediately: Numbers of the form $k/2^n$, where k and n are positive integers, do not have a unique binary representation. In other words, two different sample outcomes such as HTTTTT... and THHHH... would correspond to the same real number $1/2$ (since $1/2 = 0.0111\dots = 0.100\dots$), and the correspondence between Ω and $[0,1]$ would not, consequently, be one-to-one. We note, however, that numbers of the form $k/2^n$ constitute a denumerable set, and that there are two sample outcomes that correspond to each such number. If one, but not both, of each of these outcomes were to be removed from Ω , we would be left with a one-to-one map from a censored sample space Ω' onto $[0,1]$. Moreover, our assumption regarding individual sample points forces $P(\Omega \setminus \Omega')$ to equal zero. Thus, if a set of zero probability is thrown out from the original sample space, we may let $\Omega = [0, 1]$ and derive great satisfaction from the knowledge that this would not change the answer to any of our probability calculations.

It is possible to show, in a somewhat non-rigorous fashion (i.e., without using much measure theory), or rigorously, by introducing Lebesgue measure, that infinite coin tossing is mathematically equivalent to choosing a number randomly from the interval $[0,1]$. It can be shown, in a completely analogous way, that infinite random typewriting is equivalent to the single random choice of a number in $[0,1]$. We need of course, to consider the N -ary representation of numbers in $[0,1]$, instead of their binary expansion (where N is the number of typewriter keys). However, we shall not do so here.

Example 5. (Random Numbers) Let the random variable X denote the random choice of a number from $[0,1]$. Then

$$P(X \text{ is rational}) = \sum_{n=1}^{\infty} P(X = r_n,)$$

where r_n is the n 'th rational. Since, $P(X = r_n) = 0$ for each n , we have that

$$P(X \text{ is rational}) = 0.$$

This result may be compared with a mundane fact of “reality”: If a person, computer, pointer, or random number generator were asked to choose X , limitations of measurement accuracy (or decimal point restrictions) would systematically exclude irrational X 's, leading to the “conclusion” that

$$P(X \text{ is rational}) = 1!$$

We would like to next state a thrilling result, called *Borel's law of normal numbers* [3]: A number in $[0,1]$ is said to be *normal*, if its decimal representation has, asymptotically, an equal frequency of the digits 0 through 9:

$$\lim_{n \rightarrow \infty} \left\{ \frac{\text{the number of times } j \text{ appears in the first } n \text{ digits of its decimal expansion}}{n} \right\} = \frac{1}{10}$$

for each $j = 0, 1, 2, \dots, 9$. Borel's law states that

$$P(X \text{ is normal}) = 1,$$

which is somewhat surprising, since it is awfully hard to think of a single number that *is* normal (the number 0.012345678910111213..., obtained by writing each integer successively, is known to be normal; the proof is not trivial).

The Borel-Cantelli lemma yields several consequences that may, at first glance, seem to contradict Borel's normal number law:

Almost all the numbers in $[0,1]$ (i.e., all except some with zero Lebesgue measure) have decimal expansions that contain infinitely many chains of length 1000, say, that contain no numbers except 2,3, and 4. The nice part is, of course, that almost all of these numbers are normal *as well*, and so on.

The moral of the Borel-Cantelli lemma should, by now, be quite clear: "The realization of a truly random infinite procedure will, with probability one, contain infinitely many segments that exhibit extreme 'non-randomness', of all sizes, patterns and intensities." The Borel-Cantelli lemma is, after all, a *limit theorem* of probability, and a quote from the classic treatise of Gnedenko and Kolmogorov [9] might be in order as well: "In reality, however, the epistemological value of the theory of probability is revealed only by limit theorems. Moreover, without limit theorems, it is impossible to understand the real content of the primary concept of all our sciences—the concept of probability."

4. Conclusions and future developments

The main results of this chapter, accessible to a second-year undergraduate, are Corollaries 2.3 and 2.5. They follow from the Borel-Cantelli lemmas and Boole's inequality, respectively. Corollary 2.3 states that in an infinite sequence of keystrokes, any fixed-length "work" appears infinitely often with probability 1. Most undergraduates that the author has taught have great difficulty believing this fact, since most statistical tests, for example, are based on finite samples. Corollary 2.5 goes one step further, proving that *every* finite-length piece of work, even those yet unwritten, will *each* appear infinitely often with probability 1. The undergraduate reader will undoubtedly appreciate the "power of infinity" on reading this chapter, while graduate students will enjoy a nonpractical yet deep application of the Borel-Cantelli lemmas.

Example 4 makes a contrast between the finite situation and the infinite one. An important practical problem in this regard would be to use Poisson approximations as in [10] to find the approximate probability that a specific work occurs x times in n keystrokes and to use this process as the basis of a statistical test for randomness.

IntechOpen

IntechOpen

Author details

Anant P. Godbole
East Tennessee State University, Johnson City, USA

*Address all correspondence to: godbolea@etsu.edu

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Gamow G. One Two Three Infinity. New York: Bantam Books; 1971
- [2] Ash RB. Real Analysis and Probability. New York: Academic Press; 1972
- [3] Borel E. Les probabilités dénombrables et leurs applications arithmétiques. Rendiconti del Circolo Matematico di Palermo. 1909;**27**:247-271
- [4] Cantelli FP. Su due applicazioni di un teorema di G. Boole. Reale Accademia Nazionale dei Lincei. 1917;**26**:295-302
- [5] Balakrishnan N, Stepanov A. Generalization of the Borel-Cantelli lemma. The Mathematical Scientist. 2010;**35**:6162
- [6] Stepanov A. On strong convergence. Communications in Statistics, Theory and Methods. 2015;**44**:1615-1620
- [7] Petrov VV. A generalization of the Borel-Cantelli lemma. Statistics & Probability Letters. 2004;**67**:233-239
- [8] Billingsley P. Probability and Measure. New York: John Wiley & Sons, Inc.; 1978
- [9] Gnedenko BV, Kolmogorov AN. Limit Theorems for Sums of Independent Random Variables. Boston: Addison Wesley; 1954
- [10] Godbole A, Schaffner A. Improved Poisson approximations for word patterns. Advances in Applied Probability. 1993;**25**:334-347